Supplementary material for "Multiple tests based on a Gaussian approximation of the Unitary Events method with delayed coincidence count"

Christine Tuleau-Malot, Amel Rouis, Franck Grammont, Patricia Reynaud-Bouret

Abstract: The aim of this supplementary material is to provide the proofs of Theorems 1 and 2 and to describe more precisely the distinction between Hawkes and injection models. In both proofs, the window W is assumed to be equal to $[a, b]$ such that $T = b - a$.

A Proof of Theorem 1

Proof. Since N_1 et N_2 are independent homogeneous Poisson processes with intensity λ_1 and λ_2, one has that

$$\mathbb{E}(X) = \mathbb{E}\left(\int_a^b \int_a^b \mathbf{1}_{|x-y|\leq\delta} N_1(dx)N_2(dy)\right)$$

One can prove (see (Daley & Vere-Jones, 2003)), that

$$\mathbb{E}(X) = \lambda_1\lambda_2 \int_a^b \int_a^b \mathbf{1}_{|x-y|\leq\delta} dxdy.$$

But

$$\begin{aligned}
\int_a^b \int_a^b \mathbf{1}_{|x-y|\leq\delta} dxdy &= \int_a^b [\min(b, x+\delta) - \max(a, x-\delta)]dx \\
&= \int_a^{a+\delta} [x+\delta-a]dx + \int_{a+\delta}^{b-\delta} 2\delta dx + \int_{b-\delta}^b [b+\delta-x]dx \\
&= \frac{(a+\delta)^2 - a^2}{2} + \delta(\delta - a) + 2\delta[(b-a) - 2\delta] + \delta(b+\delta) - \frac{b^2 - (b-\delta)^2}{2} \\
&= 2\delta(b-a) - \delta^2
\end{aligned}$$

since $a+\delta \leq b-\delta$. Similarly, we can compute $\mathbb{E}(X^2)$. If $Diag = \{(x,y) \mid x = y \in [a,b]\}$ and $[a,b]^{(2)} = [a,b]^2 \backslash Diag$, one can decompose

$$\begin{aligned}
\mathbb{E}(X^2) &= \mathbb{E}\left(\int_{[a,b]^4} \mathbf{1}_{|x-y|\leq\delta}\mathbf{1}_{|t-u|\leq\delta} N_1(dx)N_1(dt)N_2(dy)N_2(du)\right) \\
&= \mathbb{E}\left(\int_{([a,b]^{(2)})^2} \mathbf{1}_{|x-y|\leq\delta}\mathbf{1}_{|t-u|\leq\delta} N_1(dx)N_1(dt)N_2(dy)N_2(du)\right) + \\
&+ \mathbb{E}\left(\int_{[a,b]^{(2)}\times Diag} \mathbf{1}_{|x-y|\leq\delta}\mathbf{1}_{|t-u|\leq\delta} N_1(dx)N_1(dt)N_2(dy)N_2(du)\right) + \\
&+ \mathbb{E}\left(\int_{Diag\times[a,b]^{(2)}} \mathbf{1}_{|x-y|\leq\delta}\mathbf{1}_{|t-u|\leq\delta} N_1(dx)N_1(dt)N_2(dy)N_2(du)\right) + \\
&+ \mathbb{E}\left(\int_{Diag^2} \mathbf{1}_{|x-y|\leq\delta}\mathbf{1}_{|t-u|\leq\delta} N_1(dx)N_1(dt)N_2(dy)N_2(du)\right)
\end{aligned}$$

This leads by classical properties of the moment measure of Poisson processes (see (Daley & Vere-Jones, 2003)) to

$$\begin{aligned}
\mathbb{E}(X^2) &= \mathbb{E}(X)^2 + \lambda_1^2\lambda_2 \int_{[a,b]^3} \mathbf{1}_{|x-y|\leq\delta}\mathbf{1}_{|t-y|\leq\delta} dxdydt + \\
&+\lambda_1\lambda_2^2 \int_{[a,b]^3} \mathbf{1}_{|x-y|\leq\delta}\mathbf{1}_{|x-u|\leq\delta} dxdydu + \lambda_1\lambda_2 \int_a^b \int_a^b \mathbf{1}_{|x-y|\leq\delta} dxdy
\end{aligned}$$

Hence

$$\mathrm{Var}(X) = \lambda_1\lambda_2[2\delta(b-a) - \delta^2] + [\lambda_1^2\lambda_2 + \lambda_1\lambda_2^2] \int_a^b \left(\int_a^b \mathbf{1}_{|x-y|\leq\delta} dy\right)^2 dx.$$

It remains to compute as before

$$\begin{aligned}
\int_a^b \left(\int_a^b \mathbf{1}_{|x-y|\leq\delta} dx\right)^2 dy &= \int_a^b [\min(b, x+\delta) - \max(a, x-\delta)]^2 dx \\
&= \int_a^{a+\delta} [x+\delta-a]^2 dx + \int_{a+\delta}^{b-\delta} [2\delta]^2 dx + \int_{b-\delta}^b [b+\delta-x]^2 dx \\
&= \frac{(a+\delta)^3 - a^3}{3} + [(a+\delta)^2 - a^2](\delta - a) + \delta(\delta-a)^2 \\
&\quad +[2\delta]^2[(b-a) - 2\delta] + \\
&\quad +\frac{b^3 - (b-\delta)^3}{3} - [b^2 - (b-\delta)^2](b+\delta) + \delta(b+\delta)^2 \\
&= 4\delta^2(b-a) - \frac{10}{3}\delta^3
\end{aligned}$$

The last part of Theorem 1 on the convergence in distribution is just a direct application of the Central Limit Theorem (Bickel & Doksum, 2000).

□

B Proof of Theorem 2

Proof. Let $X^{(m)} = \int_{[a,b]^2} \mathbf{1}_{|x-y|\leq\delta} N_1^{(m)}(dx) N_2^{(m)}(dy)$ be the coincidences count with delay δ on $[a,b]$ for the couple $(N_1^{(m)}, N_2^{(m)})$ observed during trial m.
First we apply the vectorial central limit theorem (Bickel & Doksum, 2000)

$$M^{-1/2} \sum_{m=1}^{M} \left[\begin{pmatrix} X^{(m)} \\ N_1^{(m)}([a,b]) \\ N_2^{(m)}([a,b]) \end{pmatrix} - \begin{pmatrix} \lambda_1\lambda_2[2\delta(b-a)-\delta^2] \\ \lambda_1(b-a) \\ \lambda_2(b-a) \end{pmatrix} \right] \xrightarrow{\mathcal{L}} \mathcal{N}_3(0,\Gamma),$$

where Γ is the corresponding covariance matrix, i.e.

$$\Gamma = \begin{pmatrix} \lambda_1\lambda_2[2\delta(b-a)-\delta^2] + [\lambda_1^2\lambda_2 + \lambda_1\lambda_2^2][4\delta^2(b-a) - \frac{10}{3}\delta^3] & \lambda_1\lambda_2[2\delta(b-a)-\delta^2] & \lambda_1\lambda_2[2\delta(b-a)-\delta^2] \\ \lambda_1\lambda_2[2\delta(b-a)-\delta^2] & \lambda_1(b-a) & 0 \\ \lambda_1\lambda_2[2\delta(b-a)-\delta^2] & 0 & \lambda_2(b-a) \end{pmatrix}.$$

This matrix is obtained by the previous computations, the fact that N_1 is independent of N_2 and the following fact,

$$\begin{aligned} \mathbb{E}(XN_1([a,b])) &= \mathbb{E}(\int_{[a,b]^3} \mathbf{1}_{|x-y|\leq\delta} N_1(dx)N_1(dt)N_2(dy)) \\ &= \mathbb{E}(\int_{[a,b]^{(2)}\times[a,b]} \mathbf{1}_{|x-y|\leq\delta} N_1(dx)N_1(dt)N_2(dy)) + \\ &\quad + \mathbb{E}(\int_{Diag\times[a,b]} \mathbf{1}_{|x-y|\leq\delta} N_1(dx)N_1(dt)N_2(dy)) \\ &= \lambda_1^2\lambda_2 \int_{[a,b]^3} \mathbf{1}_{|x-y|\leq\delta} dxdtdy + \lambda_1\lambda_2 \int_{[a,b]^2} \mathbf{1}_{|x-y|\leq\delta} dxdy \\ &= \mathbb{E}(X)\mathbb{E}(N_1([a,b])) + \lambda_1\lambda_2[2\delta(b-a)-\delta^2] \end{aligned}$$

Next we can rewrite

$$\begin{aligned} \sqrt{M}\,(\bar{m} - \hat{m}_0) &= \sqrt{M}\left[g\left(\frac{1}{M}\sum_{m=1}^{M} X^{(m)}, \frac{1}{M}\sum_{m=1}^{M} N_1^{(m)}([a,b]), \frac{1}{M}\sum_{m=1}^{M} N_2^{(m)}([a,b]) \right) \right. \\ &\quad \left. - g(\mathbb{E}(X^{(1)}), \lambda_1(b-a), \lambda_2(b-a)) \right], \end{aligned}$$

where $\bar{m}$ and $\hat{m}_0$ are respectively defined by Equations (11) and (14), and with $g(x,u,v) = x - uv[2\delta(b-a)-\delta^2](b-a)^{-2}$. Therefore the delta method (Bickel & Doksum, 2000) gives that

$$\sqrt{M}\,(\bar{m} - \hat{m}_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, D'\Gamma D),$$

where D is the gradient of g in $(\mathbb{E}(X), \lambda_1(b-a), \lambda_2(b-a))$. That is

$$D = \begin{pmatrix} 1 \\ -\lambda_2[2\delta(b-a)-\delta^2](b-a)^{-1} \\ -\lambda_1[2\delta(b-a)-\delta^2](b-a)^{-1} \end{pmatrix}.$$

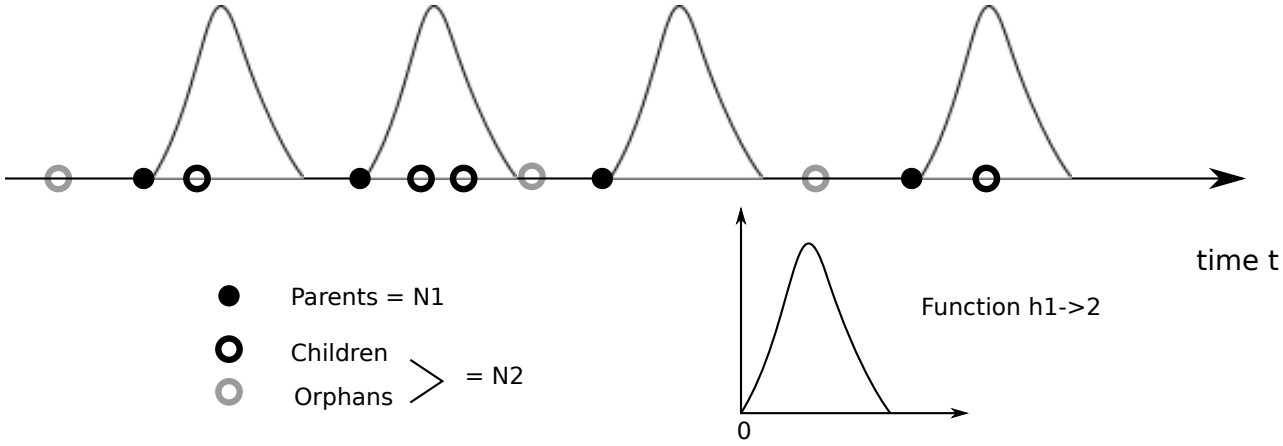

Figure 1: Interpretation of (N_1, N_2) in terms of parents, children and orphans.

We recognise that $\sigma^2 = D'\Gamma D$ and the last result is just a classical application of Slustky's lemma (Bickel & Doksum, 2000), since $\hat{v}$ converges in probability towards v. □

C Another look at the Hawkes and injection models

In some special cases, bivariate Hawkes processes, as introduced by Hawkes (1971), can be linked with injection models. Let us focus on the particular bivariate Hawkes process (N_1, N_2) whose conditional intensities are given by

$$\lambda_1(t) = \nu_1 \quad \text{and} \quad \lambda_2(t) = \nu_2 + \int_{-\infty}^{t-} h_{1\to 2}(t-u) N_1(du), \tag{1}$$

where the function $h_{1\to 2}$ has support in $\mathbb{R}_+^*$. In this case, one can see the process N_1 as the arrival times of "parents". This a homogeneous Poisson process of constant intensity ν_1. A parent, arriving at time u, is giving birth to "children" according to an inhomogeneous Poisson process of intensity $h_{1\to 2}(. - u)$. There are also orphans arriving according to a homogeneous Poisson process of constant intensity ν_2. The collection of the orphans and of all the children whatever their parents are is then forming the point process N_2 (see Figure 1).

More generally, as soon as all the interaction functions are non negative, there is a more general interpretation of the Hawkes process in terms of branching process (Hawkes & Oakes, 1974; Daley & Vere-Jones, 2003), i.e. several generations of parents and children. However as soon as there is inhibition (negative interaction functions), there is no interpretation in terms of branching processes anymore, even if thinning constructions allow to simulate the process (Ogata, 1981; Brémaud & Massoulié, 1996).

Let us go back to the case described by (1). In classical cases, one can usually assume that

$$0 < \mu := \int h_{1\to 2}(x)dx < 1.$$

This implies that the number of children per parent is in average less than 1. In particular the probability to have no children is $e^{-\mu} \simeq 1 - \mu$, as soon as μ is small enough. Note however that it is possible to have more than one child per parent even if the probability of occurrence is very small. Moreover the positions of the children with respect to the parents are random and distributed accorded to $h_{1\to 2}(.)/\mu$. In particular if the support of $h_{1\to 2}$ is really large, typically $h_{1\to 2} = \gamma \times \mathbf{1}_{[0,\delta]}$ with large δ, the position of the child can be very far away from its parent but can also be very close to the parent with equal chance.

In this respect, if $h_{1\to 2}$ tends toward $\mu\delta_0$, where δ_0 is the dirac mass in 0, a child (if it appears) is positioned at the same place as its parent. Note also that a child appears with probability close to μ as soon as μ is small enough. In this limit case, we are consequently very close to an injection model generated as follows. Take N_{B_1}, N_{B_2} and N_c three independent Bernoulli processes with firing rates $\nu_1(1-\mu)$, ν_2 and $\nu_1\mu$. Then form the injection model as described in Section 3.1 with $N'_c = N_c$ (i.e. $x = 0$), $N_1^{inj} = N_{B_1} \cup N_c$ and $N_2^{inj} = N_{B_2} \cup N'_c = N_{B_2} \cup N_c$. With this correspondence, N_1^{inj} correspond to the points (parents) in N_1, N_{B_1} correspond to the parents in N_1 that have no child, N_c to the ones with child, N_{B_2} to the orphans and N'_c to the children with parents. In this case the main difference between both models is that it may happen that several children have the same parent with the Hawkes model (even in the limit case), which is not possible with the injection model. However the probability that such things happen is so small, when μ is small that the resulting processes will be most of the time indistinguishable.

Injection models have near coincidences that are not only due to chance, by taking $x > 0$. In this case, indeed, an additional near coincidence correspond to a couple (S, S'), with $S \in N_c$, and with $S' = S + U$ in N'_c, where U is a random uniform variable on $\{-x, ..., x\} \times h$. Therefore the delay $S' - S$ is uniformly distributed in $\{-x, ..., x\} \times h$. In the Hawkes model with $h_{1\to 2} = \gamma \times \mathbf{1}_{[0,\delta]}$, an additional near coincidence correspond to a couple (S, S'), with $S \in N_1$, parent conditioned on having a child, and with S' child of S. The delay $S' - S$ is uniformly distributed in $[0, \delta]$: the child in N_2 is always after its parent and there can be several children for one parent, since the number of children per parent is a Poisson variable, even if it is rare when $\mu = \int h_{1\to 2}(x)dx = \gamma\delta$ is small.

Note finally that taking $h_{1\to 2}$ not a multiple of some $\mathbf{1}_{[0,\delta]}$, leads to other distributions for the delay $S' - S$ than the uniform one.

References

Bickel, P.J., & Doksum, K.A. (2000). *Mathematical statistics: Basic ideas and selected topics, vol. 1*. Pearson.

Brémaud, P. & Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab.*,*24*(3), 1563–1588.

Daley, D.J., & Vere-Jones, D. (2003). *An introduction to the theory of point processes, vol. I*. Springer.

Grün, S., Diesmann, M., Grammont, F., Riehle, A., & Aertsen, A.M. (1999). Detecting unitary events without discretization of time. *Journal of Neuroscience Methods*, *93*, 67–79.

Hawkes, A. G. (1971). Point spectra of some mutually exciting point processes. *Journal of the Royal Society. Series B*, *33*(3), 438–443.

Hawkes, A.G. & Oakes, D. (1974). A Cluster Process Representation of a Self-Exciting Process. *Journal of Applied Probability* , *11*(3), 493–503.

Ogata, Y. (1981). On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory*,*27*(1), 23–31.